Epidemiological Typing of Serratia marcescens Isolates by Whole-Genome Multilocus Sequence Typing

Rossen, John W. A.; Dombrecht, Jill; Vanfleteren, Diederik; De Bruyne, Katrien; van Belkum, Alex; Rosema, Sigrid; Lokate, Mariette; Bathoorn, Erik; Reuter, Sandra; Grundmann, Hajo

Published in:
Journal of Clinical Microbiology

DOI:
10.1128/JCM.01652-18

JCM Accepted Manuscript Posted Online 6 February 2019
J. Clin. Microbiol. doi:10.1128/JCM.01652-18

EPIDEMIOLOGICAL TYPING OF *SERRATIA MARCESCENS* BY WHOLE GENOME MULTI-LOCUS SEQUENCE TYPING

John W.A. Rossen[1], Jill Dombrecht[2], Diederik Vanfleteren[2], Katrien De Bruyne[2], Alex van Belkum[3,*], Sigrid Rosema[1] , Mariette Lokate[1], Erik Bathoorn[1], Sandra Reuter[4], Hajo Grundmann[4], Julia Ertel[5,6], Paul G. Higgins[5,6] and Harald Seifert[5,6]

[1] University of Groningen,
University Medical Center Groningen,
Department of Medical Microbiology and Infection Prevention, Mail Code EB80,
Hanzeplein 1, 9713 GZ Groningen, The Netherlands.
[2]bioMérieux, Data Analytics Department, Applied Maths NV,
Keistraat 120, 9830 St-Martens-Latem, Belgium.
[3]bioMérieux, Data Analytics Department,
3 Route de Port Michaud, 38390 La Balme Les Grottes, France.
[4]Medical Center – University of Freiburg, Faculty of Medicine,
Institute for Infection Prevention and Hospital Epidemiology
Breisacher Str. 115 B, 79106 Freiburg, Germany
[5]Institute for Medical Microbiology, Immunology and Hygiene,
University of Cologne, Goldenfelsstrasse 19-21, 50935 Köln, Germany
[6]German Centre for Infection Research (DZIF),
Partner Site Bonn-Cologne, Germany

[*]Communicating author: bioMerieux Data Analytics Department, 3 Route de Port Michaud, La Balme Les Grottes, France
e-mail alex.vanbelkum@biomerieux.com
phone +33609487905

Key words: *Serratia marcescens* – outbreak management – BioNumerics™ – neonatal intensive care – molecular typing - whole genome sequencing (WGS) – whole genome Multi Locus Sequence Typing (wgMLST).

ABSTRACT

Serratia marcescens is an opportunistic bacterial pathogen. It is notorious for its increasing antimicrobial resistance and its potential to cause outbreaks of colonization and infections predominantly in neonatal intensive care units (NICUs). There, its spread requires rapid infection control response. In order to understand its spread, detailed molecular typing is key. We present a whole genome multi-locus sequence typing (wgMLST) method for *S. marcescens*. Using a set of 299 publicly available whole genome sequences (WGS) we developed an initial wgMLST system consisting of 9377 gene loci. This included 1455 loci occurring in all reference genomes and 7922 accessory loci. This closed system was validated using three geographically diverse collections of *S. marcescens* consisting of 111 clinical isolates implicated in nosocomial dissemination events in three hospitals. The validation procedure showed a full match between epidemiological data and the wgMLST analyses. We set the cut-off value for epidemiological (non-)relatedness at 20 different alleles, although for the majority of outbreak-clustered isolates this difference was limited to 4 alleles. This shows that the wgMLST system for *S. marcescens* provides prospects of successful future monitoring for the epidemiological containment of this opportunistic pathogen.

Journal of Clinical Microbiology JCM Journal of Clinical Microbiology

INTRODUCTION

The new Gold Standard in microbial epidemiology is genome sequencing. The use of whole genome (draft) sequences (WGS) to compare bacterial isolates in detail, and to delineate their spread, is based on either the detection of single nucleotide variants or polymorphisms (SNVs and SNPs) or on the assessment of overall gene content including allelic differences between strains by whole genome multi-locus sequence typing (wgMLST) (1-4). Both methods have their advantages and disadvantages. Where SNP analysis may have a higher intrinsic discriminatory power (since it covers coding and non-coding regions) and better resolves the ancestral relationship between lineages, wgMLST usually provides a more stable, generically applicable system, with results that are easier to translate into relevant epidemiological differences between isolates. wgMLST schemes have been developed for a multitude of microbial organisms, with the main driver being the development of a universal “typing language” (5-7). This will facilitate the monitoring of local, institutional spread of certain pathogens but will also extend into regional, national, international, and possibly even global monitoring for the dissemination of given bacterial strain types (8-10). This will aid communication in international public health management and should in the end lead to early recognition of the emergence and spread of pathogenic microbial strains. Furthermore, this is of importance in the current era of multi-drug resistant bacteria and their global dispersal promoted by human travelling, international patient transfer, nosocomial transmission, and excessive use of antimicrobials.

Serratia marcescens is a bacterial pathogen for which no wgMLST scheme has been defined yet. *S. marcescens* is notorious for its pathogenicity in plants (11) but also in preterm neonates (12,13). Therefore, setting up a robust epidemiological wgMLST typing scheme is essential for monitoring and interrupting outbreaks in neonatal intensive care units (NICU) as well as other medical settings. In addition, *S. marcescens* is capable of efficiently acquiring multiple resistance determinants (that are unreliable epidemiological markers) which adds to its clinical relevance (14-18). We have developed a proprietary wgMLST toolbox for *S. marcescens* based on publicly available WGS data. We have validated the scheme using epidemiologically related isolates collected during recent outbreaks of colonization and infection in NICUs in both Dutch and German teaching hospitals.

Journal of Clinical Microbiology

MATERIALS AND METHODS

Strains: Clinical *S. marcescens* isolates were obtained from three different institutions in Groningen (The Netherlands; n=41), Cologne (Germany; n=19) and Freiburg (Germany; n=51), respectively.

The 41 isolates from the University Medical Center Groningen were obtained between 2014 and 2017 from 38 patients of which 4 were adults in non-pediatric wards (2 in cardiology, 1 in orthopedics and 1 in obstetrics), 2 were from children > 12 year in the pediatric ICU (PICU), 1 from a child > 18 months and the others from children < 6 months either on the pediatric special care unit (n=1), the pediatric general surgery ward (n=2), the PICU (n=2), or the NICU (n=26). From three patients two isolates were sequenced. In one case, in addition to a positive culture from a rectal swab of the patient also an isolate was cultured from the intravenous line, but this isolate appeared to be a *S. liquefaciens,* originally misidentified as *S. marcescens* by conventional diagnostic methods. All other isolates were cultured from patients in the NICU using growth-based microbiology technology (see Figure 1 for additional details on strain origin). The 19 isolates from Cologne were isolated between 2014 and 2017 and all originate from NICUs, PICUs and general wards. The age of the patients varied between 4 days and 11 months. The collection of isolates consisted of 5 epidemiologically related transmission clusters and 2 singleton isolates (see Figure 2 for additional details). The 51 isolates from Freiburg mostly originated from the local NICU (n=39) with patient age varying between 0 and 12 weeks. Seven environmental isolates were included for comparative reasons and to gauge the relevance of environmental spread. For several patients (A to H, n=8) multiple isolates were included in order to define basic levels of intra-patient variability of *S. marcescens* (see Figure 3 for additional details).

Isolates were either directly processed or stored at -80^{o}C in glycerol-containing media until culture for DNA isolation and genome sequencing. In addition to the WGS data, clinical and epidemiological data were included. Metadata included, but were not limited to, isolation dates, outbreak associations, patients' gender and age, type (and outcome) of infections, specimen types submitted for microbiological analyses, location of the ward and whether local typing data obtained previously were available.

DNA isolation: DNA was extracted using the Ultraclean Microbial DNA isolation kit (MoBio Laboratories, Carlsbad, CA, USA) or the MagAttract HMW DNA Isolation kit, in both cases following the manufacturer's instructions (Qiagen, Hilden, Germany) and quantified using a NanoDrop 2000c spectrophotometer (Thermo Fisher Scientific Inc, Waltham, MA, USA) and/or the Qubit dsDNA HS assay (Thermo Fisher Scientific GmbH, Schwerte, Germany).

Genome sequencing: DNA libraries were prepared using the Nextera XT library preparation kit and the Nextera XT v2 index kit (Illumina, San Diego, CA, USA). The library was sequenced on a MiSeq, using the reagent kit v2 generating 250-bp paired-end reads. Supplementary Tables 1A to 1C disclose the quality parameters for the sequences determined. All WGS included met with the required quality criteria and all primary sequences were deposited in the public domain (ENA project numbers PRJEB28358 and PRJEB28681).

Development of the wgMLST scheme: A scheme for wgMLST of *S. marcescens* was developed using publicly available WGS data for this species (June 2017), and will be made commercially available through a plugin in BioNumerics™ (Applied Maths NV, St-Martens-Latem, Belgium). The scheme is intended to facilitate detection of subtype- or outbreak-specific markers. Using a selection of 299 annotated, publicly available reference genomes which were assumed to capture the diversity within *S. marcescens*, a pan-genomic scheme with high discriminatory power was developed (see Supplementary Table 2 for a list of all WGS included). Starting from the reference genomes, our scheme creation procedure uses a sampling-based multi-reciprocal BLAST procedure to determine those sets of alleles that make up the stable loci in the pan-genome. A per-locus allele assessment procedure then determines the central prototype allele, and thus the definition of the locus. The wgMLST scheme for *S. marcescens* was tested, validated and approved by epidemiological and microbiological analyses using information on the strain collections from Groningen, Cologne and Freiburg.

Bioinformatic analyses: De novo genome assembly for all WGS was performed using SPAdes 3.7.1. All de novo calculations were run on the cloud-based calculation engine that comes with BioNumerics™ 7.6.3. wgMLST analysis was also performed using the BioNumerics™ cloud-based calculation engine. Alleles were identified by both an assembly-free k-mer based approach using the raw reads and

Journal of Clinical Microbiology

an assembly-based BLAST approach. Identification was done against the *S. marcescens* wgMLST database in BioNumerics™. Categorical coefficients were used for defining similarity levels and Unweighted Pair Group Method with Arithmetic Mean (UPGMA) was used as clustering algorithm. Minimum spanning trees (MST) were constructed using the wgMLST allelic profiles as input data. The size of the nodes was chosen proportional to the number of isolates in the nodes (i.e. isolates with the same allelic profiles). Branch lengths reflect the number of allele differences between the isolates in the connected nodes.

RESULTS

A new system for wgMLST for *S. marcescens*: In total, 299 reference genome sequences were included while building the wgMLST scheme. These displayed a conformity between 85% and 97% after constructing the scheme and showed an average of 95% global coverage of the included loci. The scheme was validated in August 2017 on the basis of 373 sequence read archives (SRA), which included all Illumina data sets publicly available as of 28 August 2017. In this way, a total of 9,377 loci were added to the scheme, including 1455 loci which were present in all references and 7922 accessory loci. The wgMLST scheme had high discriminatory power and allowed for the detection of markers specific for *S. marcescens* subtypes or outbreak strains, thus enabling powerful classification and outbreak definition (see Figure 4C). The two allele detection procedures (either assembly-based or assembly-free) performed fast and reliable allele calling for cluster detection. Figure 4A indicates the diversity within the reference genome set, and provides an overview of the number of clusters as function of the similarity cutoff value, indicating the presence of both distant and highly related isolates in the reference set of 299 strains. Figure 4B depicts the number of pairwise allelic differences and the frequency of their occurrence peaking at about 4000 allelic differences given the current wgMLST scheme complexity. Figure 4C shows a global perspective of the genomic diversity among the references used to build the wgMLST scheme, where all circles identify distinct wgMLST types (as also semi-quantified by the number of allelic differences quantified on the branches) and the colored blocks identify isolates of more closely related and sometimes indistinguishable genomic sequences. This confirms our assumption that the genome sequences obtained from the public

domain show significant levels of diversity allowing them to serve as reference of genomic variability. Overall, the quality parameters indicate that the scheme covers the diversity within the species and provides sufficient resolving power for distinguishing even closely related bacterial isolates. Finally, it seems that the population structure of *S. marcescens* is largely genetically diverse with many singletons present. However, there seem to be indications for the successful expansion of clones (colored circles, Figure 4C).

Strain characteristics and outbreak features: It has to be stated that only one patient died as a consequence of *S. marcescens* colonization/infection. Also, presence was mostly due to colonization and real infection was only apparent in a limited number of cases (Groningen 9 of 38 patients (24%); Cologne 2/16 (13%); Freiburg 6/23 (26%) (one sample of unknown origin)). Overall, 22% patients had an infection.

***Groningen outbreak analyses*:** Forty-one clinical isolates were obtained from 38 patients in the University Medical Center Groningen (UMCG). The wgMLST analysis detected a small cluster of related isolates: five isolates obtained from three patients in May-June in 2015 (cluster 0003 in Figure 1). From one patient two isolates from the rectal swab appeared to be 100% wgMLST identical and from the other patient the isolate found in the blood was identical to the one found in the rectal swab. In addition, a larger cluster was found containing isolates, all from different patients, from a protracted outbreak in August-November 2014 (cluster 0005 in Figure 1). The single invasive isolate that was isolated during this episode was indistinguishable from the other isolates. In addition, four suspected cases of single transmission events involving two patients were confirmed as well (clusters 0001, 0002, 0004, 0006 and 0007 in Figure 1). Hence, the clustering aligns very well with the prior epidemiological scenarios. The 0002 cluster contained two separate isolates from the same patient, showing the reproducibility of the method. All isolates contained the aminoglycoside resistance-associated gene *aac*(6')-I-C and about half of them contained the tetracycline resistance determinant Tet (41). A single multi-resistant isolate was cultured from the synovial fluid of an elderly female nursed at the orthopedics department. The origin of this strain is not clear.

***Cologne outbreak analyses*:** wgMLST analysis of the 19 isolates from the Cologne University hospital correctly defined the anticipated clustering and identified

Journal of Clinical Microbiology

two main outbreak clusters and three cases where inter-patient transfer was already suspected (Cologne-1 to Cologne-5). The two singleton isolates were separated from all of the other isolates. Figure 2 summarizes the overall data and sketches the outbreak scenarios also showing that all related isolates were 100% identical at the wgMLST level. One of the singleton isolates contained at least 8 different resistance genes.

***Freiburg outbreak analyses*:** The collection of isolates derived from the laboratory in the Freiburg University hospital contained 47 out of 51 isolates that were nearly indistinguishable by wgMLST (Figure 3, green boxes), indicating a local outbreak which occurred in October and November 2015 involving 19 patients and 7 environmental isolates. Additionally, two isolates were identified (red boxes, Figure 3) that were not distinguished by wgMLST, reflecting a single, known transmission event of a different strain type outside the NICU. Most of the outbreak isolates were considered to represent colonization rather than infection or bacteremia (16/19 patients). All serial isolates obtained from individual patients were identical at the wgMLST level. Only in case of patients F and H small differences were documented but within the boundaries of the epidemiological cut-off value. Finally, the environmental isolates all fell within the same outbreak category.

***Minimum spanning trees*:** Figure 5 displays the minimum spanning trees for the three studies and there is good concordance with the UPGMA trees in Figures 1-3. The number of allele differences ranged between 0 and 4 for the epidemiologically defined strain clusters with two exceptions. There is only a single strain in the Freiburg cluster that differs by 18 alleles from its counterparts. This suggests that a cut-off value of <20 alleles would represent a conservative but useful estimate for transmission-related isolates, also given the significantly higher genetic distance between the non-related *S. marcescens* isolates. Figure 6 once more displays the robustness of the wgMLST scheme since while including all WGS entries in the database, still the strain clusters identified above remain unchanged.

DISCUSSION

S. marcescens is a nosocomial pathogen of clinical importance and both species identification and antimicrobial susceptibility testing are well covered in routine diagnostic clinical microbiology laboratories. However, epidemiological typing of *S.*

Journal of Clinical Microbiology

marcescens is less developed, and for this reason we developed a wgMLST scheme. The system allowed for the adequate recognition of clonally related organisms and it allowed for the detection of outbreak events. At the level of wgMLST the number of changes between the most closely related isolates were less than twenty alleles (given the time frame during which our outbreak related strains were captured), although a significant fraction of the closely related genomes only differed by 0-4 alleles. This latter level of resolution does not allow for detailed epidemiological tracing of spread from one patient to the other given the apparently low number of changes associated with such transfers. We performed a limited number of wgSNP analyses and, surprisingly, for the ten related isolates from Groningen, this did not increase the resolution. The number of SNPs encountered between the ten isolates ranged from zero to five, in the same range as the wgMLST variation and insufficient to decipher transmission of strains between patients (data not shown). Of note, a recent cgMLST study for *Brucella melitensis* revealed similar findings: epidemiological cut off values for non-variance were defined as <6 loci for wgMLST and <7 loci for wgSNP analyses, similar to what we document here for *S. marcescens* (19).

Next generation sequencing (NGS) is becoming very popular in clinical microbiology (20,21), but wgMLST for *S. marcescens* has not yet been described. WGS has been used to study *S. marcescens* virulence after wound infections and infection of snake bites (22). As well, WGS has been used to study the national dissemination of drug resistance elements and plasmids in *S. marcescens* throughout Germany, although only a limited number of isolates were subjected to NGS (16). The same authors also demonstrated that *S. marcescens* genomes can be used to generate genomic catalogues of antibiotic resistance genes. Relevant to our current clinical study is the work done by Iguchi et al (23). These authors defined the WGS for two selected *S. marcescens* strains. Their analysis revealed a degree of genetic heterogeneity that our current study exploited. Iguchi et al (23) already tried to define core and variable genes and used ways for defining genetic distance between *S. marcescens* isolates. Most recently, Martineau and colleagues (24) used WGS to elucidate transmission patterns in a NICU in Montreal, Canada. WGS for ten clinical isolates were instrumental in resolving *S. marcescens* routes of spread in this setting. We were able to confirm their data using our wgMLST scheme (results not shown). In the

examples brought forward by Martineau et al, a single outbreak was analyzed, where we have now taken the method to a higher level including the development of a dedicated wgMLST WGS database and an informatics tool for the semi-automated analysis of potential outbreak scenarios. With turnaround calculation times of less than 30 minutes per sample and simultaneous processing of up to 24 samples, high-powered wgMLST performance is guaranteed. Using BioNumerics™ and a cloud-based calculation engine, it provides a high-throughput environment that enables a fast and simple outbreak analysis of WGS data for *S. marcescens*. The calculation engine's quality-controlled *de novo* assembly possibilities allow for rapid, push-button assembly of WGS data without the need of local computing power. In short, even high resolution typing needs optimal epidemiological data and cannot stand on its own. Although we here focus on patients in NICUs it should be emphasized that genomic typing of *S. marcescens* will have wider implications as these bacteria infect other risk groups as well (25,26). We acknowledge the fact that we are not disclosing the precise methodology used for wgMLST scheme development since this module will become available only in combination with BioNumerics™.

In conclusion, all laboratory-run typing methods, wgMLST included, are valuable in the context of hospital-wide screening for pathogens but also for analyses of random clinical isolates (27,28). wgMLST for *S. marcescens* has here been demonstrated to be a promising epidemiological typing support tool. In combination with tools for deciphering a genomic antibiogram and the presence of virulence genes, WGS by NGS may help trace and follow outbreaks, understand the acquisition and spread of resistance factors and explain the disease invoking potential for this not-to-be-underestimated human pathogen.

ACKNOWLEDGEMENTS

This work was done in collaboration with the European Society of Clinical Microbiology and Infectious Diseases (ESCMID) Study Group on Genomic and Molecular Diagnostics (ESGMD), and the ESCMID Study Group on Epidemiological Markers (ESGEM), Basel, Switzerland.

TRANSPARENCY DECLARATION

Journal of Clinical Microbiology

Alex van Belkum, Jill Dombrecht, Diederik Vanfleteren and Katrien De Bruyne are employees of bioMérieux, a company designing, developing and selling infectious disease diagnostics and hence have a business implication in this work. John Rossen consults for IDbyDNA. All other authors declare no conflicts of interest and have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed. No external financial support was provided for the studies presented herein.

LITERATURE REFERENCES

1. Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl JM, Laurent F, Grundmann H, Friedrich AW; ESCMID Study Group of Epidemiological Markers (ESGEM). 2013. Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Euro Surveill 18:20380.

2. Goering RV, Köck R, Grundmann H, Werner G, Friedrich AW; ESCMID Study Group for Epidemiological Markers (ESGEM). 2013. From theory to practice: molecular strain typing for the clinical and public health setting. Euro Surveill 18:20383.

3. Higgins PG, Prior K, Harmsen D, Seifert H. 2017. Development and evaluation of a core genome multilocus typing scheme for whole-genome sequence-based typing of *Acinetobacter baumannii*. PLoS One 12:e0179228.

4. Rossen JWA, Friedrich AW, Moran-Gilad J; ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD). 2018. Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. Clin Microbiol Infect 24:355-360. doi: 10.1016/j.cmi.2017.11.001.

5. Revez J, Zhang J, Schott T, Kivistö R, Rossi M, Hänninen ML. 2014. Genomic variation between *Campylobacter jejuni* isolates associated with milk-borne-disease outbreaks. J Clin Microbiol 52:2782-6. doi: 10.1128/JCM.00931-14.

6. Jolley KA, Maiden MC. 2014. Using multilocus sequence typing to study bacterial variation: prospects in the genomic era. Future Microbiol 9:623-30. doi: 10.2217/fmb.14.24.

7. Kingry LC, Rowe LA, Respicio-Kingry LB, Beard CB, Schriefer ME, Petersen JM. 2016. Whole genome multilocus sequence typing as an epidemiologic tool for

Yersinia pestis. Diagn Microbiol Infect Dis 84:275-80. doi: 10.1016/j.diagmicrobio.2015.12.003.

8. Kluytmans-van den Bergh MF, Rossen JW, Bruijning-Verhagen PC, Bonten MJ, Friedrich AW, Vandenbroucke-Grauls CM, Willems RJ, Kluytmans JA. 2016. Whole-Genome Multilocus Sequence Typing of Extended-Spectrum-Beta-Lactamase-Producing Enterobacteriaceae. J Clin Microbiol 54:2919-27.

9. Fiebig L, Kohl TA, Popovici O, Mühlenfeld M, Indra A, Homorodean D, Chiotan D, Richter E, Rüsch-Gerdes S, Schmidgruber B, Beckert P, Hauer B, Niemann S, Allerberger F, Haas W. 2017. A joint cross-border investigation of a cluster of multidrug-resistant tuberculosis in Austria, Romania and Germany in 2014 using classic, genotyping and whole genome sequencing methods: lessons learnt. Euro Surveill 22: 30439. doi: 10.2807/1560-7917.

10. Nadon C, Van Walle I, Gerner-Smidt P, Campos J, Chinen I, Concepcion-Acevedo J, Gilpin B, Smith AM, Man Kam K, Perez E, Trees E, Kubota K, Takkinen J, Nielsen EM, Carleton H; FWD-NEXT Expert Panel. 2017. PulseNet International: Vision for the implementation of whole genome sequencing (WGS) for global food-borne disease surveillance. Euro Surveill pii:30544. doi: 10.2807/1560-7917.ES.2017.22.23.30544.

11. Besler KR, Little EL. 2017. Diversity of *Serratia marcescens* strains associated with cucurbit yellow vine disease in Georgia. Plant Dis 101:129-36. doi: 10.1094/PDIS-05-16-0618-RE.

12. Mahlen SD. 2011. Serratia infections: from military experiments to current practice. Clin Microbiol Rev 24:755-91. doi: 10.1128/CMR.00017-11.

13. Dramowski A, Aucamp M, Bekker A, Mehtar S. 2017. Infectious disease exposures and outbreaks at a South African neonatal unit with review of neonatal outbreak epidemiology in Africa. Int J Infect Dis 57:79-85. doi: 10.1016/j.ijid.2017.01.026.

14. Batah R, Loucif L, Olaitan AO, Boutefnouchet N, Allag H, Rolain JM. 2015. Outbreak of *Serratia marcescens* Coproducing ArmA and CTX-M-15 Mediated High Levels of Resistance to Aminoglycoside and Extended-Spectrum Beta-Lactamases, Algeria. Microb Drug Resist 21:470-6. doi: 10.1089/mdr.2014.0240.

15. Gruber TM, Göttig S, Mark L, Christ S, Kempf VA, Wichelhaus TA, Hamprecht A. 2015. Pathogenicity of pan-drug-resistant *Serratia marcescens* harboring blaNDM-1. J Antimicrob Chemother 70:1026-30. doi: 10.1093/jac/dku482.

16. Wendel AF, Kaase M, Autenrieth IB, Peter S, Oberhettinger P, Rieber H, Pfeffer K, MacKenzie CR, Willmann M. 2017. Protracted Regional Dissemination of GIM-1-Producing *Serratia marcescens* in Western Germany. Antimicrob Agents Chemother 23;61. doi:10.1128/AAC.01880-16.

17. Mataseje LF, Boyd DA, Delport J, Hoang L, Imperial M, Lefebvre B, Kuhn M, Van Caeseele P, Willey BM, Mulvey MR. 2014. *Serratia marcescens* harbouring SME-type class A carbapenemases in Canada and the presence of blaSME on a novel genomic island, SmarGI1-1. J Antimicrob Chemother 69:1825-9. doi: 10.1093/jac/dku040.

18. Rodríguez C, Brengi S, Cáceres MA, Mochi S, Viñas MR, Rizza CA, Merletti G, Bru E, Assa JD, Raya RR, Centrón D. 2018. Successful management with fosfomycin/ceftazidime of an infection caused by multiple highly related subtypes of MDR and XDR KPC-producing *Serratia marcescens*. Int J Antimicrob Agents pii:S0924-8579(18)30219-X. doi: 10.1016/j.ijantimicag.2018.07.020.

19. Janowicz A, De Massis F, Ancora M, Cammà C, Patavino C, Battisti A, Prior K, Harmsen D, Scholz H, Zilli K, Sacchini L, Di Giannatale E, Garofolo G. 2018. Core Genome Multilocus Sequence Typing and Single Nucleotide Polymorphism Analysis in the epidemiology of *Brucella melitensis* infections. J Clin Microbiol 56:e00517-18. doi: 10.1128/JCM.00517-18.

20. Westblade LF, van Belkum A, Grundhoff A, Weinstock GM, Pamer EG, Pallen MJ, Dunne WM Jr. 2016. Role of clinicogenomics in infectious disease diagnostics and public health microbiology. J Clin Microbiol 54:1686-93. doi: 10.1128/JCM.02664-15.

21. Dunne WM Jr, Westblade LF, Ford B. 2012. Next-generation and whole-genome sequencing in the diagnostic clinical microbiology laboratory. Eur J Clin Microbiol Infect Dis 31:1719-26. doi: 10.1007/s10096-012-1641-7.

22. Huang YT, Cheng JF, Liu YT, Mao YC, Wu MS, Liu PY. 2018. Genome-based analysis of virulence determinants of a *Serratia marcescens* strain from soft tissues following a snake bite. Future Microbiol 13:331-43. doi: 10.2217/fmb-2017-0202.

Journal of Clinical Microbiology

23. Iguchi A, Nagaya Y, Pradel E, Ooka T, Ogura Y, Katsura K, Kurokawa K, Oshima K, Hattori M, Parkhill J, Sebaihia M, Coulthurst SJ, Gotoh N, Thomson NR, Ewbank JJ, Hayashi T. 2014. Genome evolution and plasticity of *Serratia marcescens*, an important multidrug-resistant nosocomial pathogen. Genome Biol Evol 6:2096-110. doi: 10.1093/gbe/evu160.

24. Martineau C, Li X, Lalancette C, Perreault T, Fournier E, Tremblay J, Gonzales M, Yergeau É, Quach C. 2018. *Serratia marcescens* outbreak in a neonatal intensive care unit (NICU): new insights from next-generation sequencing applications. J Clin Microbiol JCM.00235-18. doi: 10.1128/JCM.00235-18.

25. Leng P, Huang WL, He T, Wang YZ, Zhang HN. 2015. Outbreak of *Serratia marcescens* postoperative infection traced to barbers and razors. J Hosp Infect 89:46-50. doi: 10.1016/j.jhin.2014.09.013.

26. Us E, Kutlu HH, Tekeli A, Ocal D, Cirpan S, Memikoglu KO. 2017. Wound and soft tissue infections of *Serratia marcescens* in patients receiving wound care: A health care-associated outbreak. Am J Infect Control 45(4):443-7. doi: 10.1016/j.ajic.2016.11.015.

27. Dawczynski K, Proquitté H, Roedel J, Edel B, Pfeifer Y, Hoyer H, Dobermann H, Hagel S, Pletz MW. 2016. Intensified colonization screening according to the recommendations of the German Commission for Hospital Hygiene and Infectious Diseases Prevention (KRINKO): identification and containment of a *Serratia marcescens* outbreak in the neonatal intensive care unit, Jena, Germany, 2013-2014. Infection 44:739-46.

28. Åttman E, Korhonen P, Tammela O, Vuento R, Aittoniemi J, Syrjänen J, Mattila E, Österblad M, Huttunen R. 2018. A *Serratia marcescens* outbreak in a neonatal intensive care unit was successfully managed by rapid hospital hygiene interventions and screening. Acta Paediatr 107:425-9. doi: 10.1111/apa.14132.

LEGENDS TO THE FIGURES

Figure 1 UPGMA tree of the pan-genomic allelic profiles (n=25) derived for *S. marcescens* isolates from the University Medical Center Groningen, The Netherlands. Outbreaks and transfer events identified prior to our study (0001-0007) are

Journal of Clinical Microbiology

highlighted by color, with relevant microbiological, host-associated and environmental metadata displayed to the right. The UPGMA tree which was built using a similarity coefficient based on categorical values expressed as a percentage. Strain UMCG-029, located at the bottom of the tree, represents *S. liquefaciens,* a species only sharing about 2900 loci with the *S. marcescens* wgMLST scheme, as opposed to 4300 loci that are typically detected in *S. marcescens*.

Figure 2 UPGMA tree of the pan-genomic allelic profiles (n=7) derived for *S. marcescens* isolates from the Institute for Medical Microbiology, Immunology and Hygiene at the University of Cologne, Germany. Outbreaks and transfer events (Cologne-1 to Cologne-5) identified prior to our study are highlighted by color, with relevant microbiological, host-associated and environmental metadata displayed to the right. The UPGMA tree which was built using a similarity coefficient based on categorical values expressed as a percentage. Isolates originating from inanimate surfaces are highlighted in blue.

Figure 3 UPGMA tree of the pan-genomic allelic profiles (n=4) derived for *S. marcescens* isolates from the University Hospital of Freiburg, Germany. A single major outbreak event generated all strains except four (red and non-boxed). Relevant microbiological, host-associated and environmental metadata are displayed to the right. The UPGMA tree was built using a similarity coefficient based on categorical values expressed as a percentage. Note that in this case multiple isolates were included for 8 different individuals. Isolates originating from inanimate surfaces are highlighted in blue.

Figure 4 Review of quality parameters for the *S. marcescens* specific whole genome sequences used to construct the wgMLST reference database.

Figure 4A Correlation between number of clusters and similarity cutoff values for the founding *S. marcescens* wgMLST database. The cluster index was based on the average number of alleles being different between closely related strain pairs. The analysis was performed using all WGS listed in Supplementary Table 2.

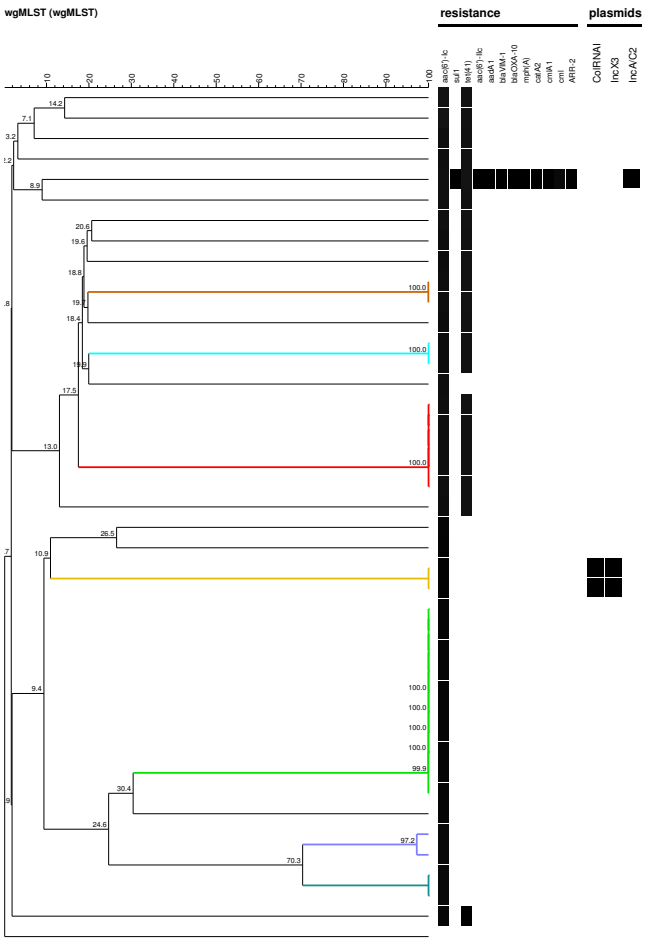
Figure 4B Correlation between the numbers of pairwise allelic differences and their frequency of occurrence.

Figure 4C Minimum spanning tree based on the pan-genomic allelic profiles of 299 *S. marcescens* isolates, representing the reference set used to create the wgMLST database. Colors highlight closely related isolates, numbers of allelic differences are indicated on the lines connecting the various types.

Figure 5 Minimum spanning trees for the *S. marcescens* isolates from Groningen, Cologne and Freiburg built from the pan-genomic allelic profiles. Colors of the circles identify the epidemiological clusters and cases of transmission. Figures on the axes identify the numbers of allelic differences between the connected isolates. Circle size is associate with the number of isolates per type. The figure implies that there are no clusters extending across hospitals. Color codes are specific for the three different panels and should not be compared between panels.

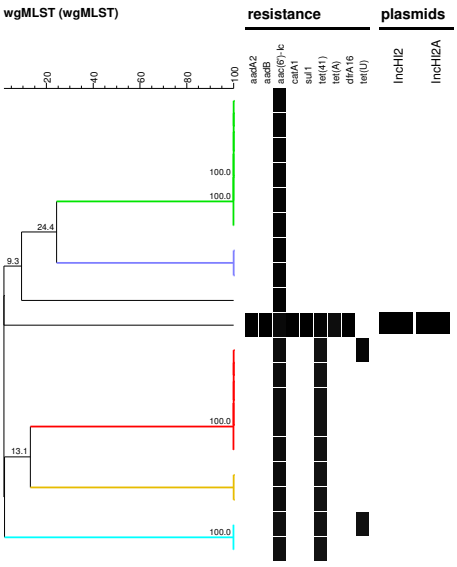
Figure 6 Overall genomic population structure of *S. marcescens* based on a combined analysis of our epidemiologically related isolates and the reference genomes that were used to construct the wgMLST scheme. Note the extended number of singletons and the occurrence of epidemic clones seemingly originating from several of such singletons. Green bullets represent isolates from Groningen, red ones the isolates from Cologne and blue ones identify the isolates from Freiburg.

Figure 1



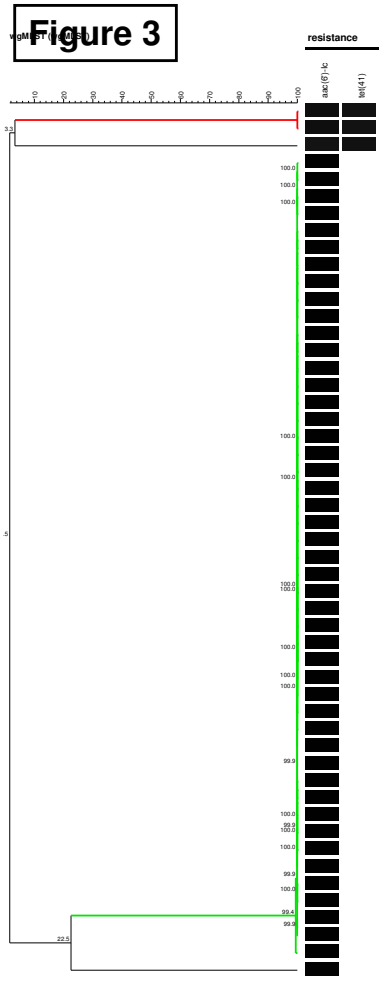
Cluster	Isolate ID	Gender	Type of Specimen	Ward	Isolation Date	Age (years)
	UMCG-041	F	Faeces	Neonatal ICU	2017-10-30	0
	UMCG-042	F	Faeces	Neonatal ICU	2017-10-30	0
	UMCG-025	M	Faeces	Neonatal ICU	2015-07-13	0
	UMCG-019	M	Faeces	Neonatal ICU	2015-01-19	0
	UMCG-031	F	Synovial fluid	Orthopedics	2016-04-05	77
	UMCG-037	M	Faeces	Childrens general surgery	2017-07-04	0
	UMCG-026	M	Faeces	Neonatal ICU	2015-07-13	0
	UMCG-030	F	Rectal	Neonatal ICU	2015-12-28	0
	UMCG-018	F	Pus thorax	Childrens ICU	2014-12-14	15
0001	UMCG-002	M	Faeces	Neonatal ICU	2014-08-18	0
0001	UMCG-005	M	Rectal	Neonatal ICU	2014-09-03	0
	UMCG-011	F	Urine	Obstetrics	2014-10-14	37
0002	UMCG-027	M	Faeces	Neonatal ICU	2015-11-17	0
0002	UMCG-028	M	Faeces	Neonatal ICU	2015-11-17	0
	UMCG-039	F	Faeces	Neonatal ICU	2017-10-23	0
0003	UMCG-020	F	Faeces	Neonatal ICU	2015-05-26	0
0003	UMCG-021	F	Faeces	Neonatal ICU	2015-05-26	0
0003	UMCG-022	M	Faeces	Neonatal ICU	2015-05-26	0
0003	UMCG-024	M	Blood culture	Neonatal ICU	2015-06-10	0
0003	UMCG-023	M	Faeces	Neonatal ICU	2015-06-08	0
	UMCG-040	M	Faeces	Neonatal ICU	2017-10-26	0
	UMCG-014	F	Faeces	Neonatal ICU	2014-10-30	0
	UMCG-034	M	Faeces	Neonatal ICU	2017-06-06	0
0004	UMCG-016	M	Sputum	Childrens ICU	2014-12-09	15
0004	UMCG-017	F	Sputum	Childrens special care	2014-12-15	0
0005	UMCG-004	M	Blood culturectal	Neonatal ICU	2014-08-29	0
0005	UMCG-006	F	Faeces	Neonatal ICU	2014-09-15	0
0005	UMCG-007	F	Faeces	Neonatal ICU	2014-09-22	0
0005	UMCG-009	M	Faeces	Neonatal ICU	2014-09-25	0
0005	UMCG-010	F	Urine	Childrens general surgery	2014-10-11	0
0005	UMCG-003	M	Sputum	Neonatal ICU	2014-08-25	0
0005	UMCG-008	M	Sputum	Childrens ICU	2014-09-23	0
0005	UMCG-012	F	Sputum	Childrens ICU	2014-10-16	0
0005	UMCG-015	F	Faeces	Neonatal ICU	2014-11-24	0
0005	UMCG-013	M	Sputum	Neonatal ICU	2014-10-28	0
	UMCG-001	F	Blood culturectal	Childrens special care	2014-06-27	1
0006	UMCG-032	M	Blood culturectal	Cardiology	2016-06-21	38
0006	UMCG-038	M	Faeces	Neonatal ICU	2017-07-18	0
0007	UMCG-035	M	Faeces	Neonatal ICU	2017-06-29	0
0007	UMCG-036	F	Faeces	Neonatal ICU	2017-06-29	0
	UMCG-033	M	Blood culturectal	Cardiology	2016-06-22	47
	UMCG-029	F	IV line	Neonatal ICU	2015-12-17	0

Figure 2



Cluster	Outbreak Info	Isolate ID	Gender	Type of Specimen	Ward	Isolation Date	Age (years)
0001	Cologne-5	AML_0403	M	Nose/Throat swab	Neonatal ward	2017-10-12	0
0001	Cologne-5	AML_0404	F	Nose/Throat swab	Neonatal ward	2017-10-05	0
0001	Cologne-5	AML_0402	M	Umbilical swab	Neonatal ward	2017-09-17	0
0001	Cologne-5	AML_0401	M	Nose/Throat swab	Neonatal ward	2017-11-27	0
0001	Cologne-5	AML_0400	F	Nose/Throat swab	Neonatal ward	2017-10-19	0
0001	Cologne-5	AML_0405	F	Nose/Throat swab	Neonatal ward	2017-12-07	0
0003	Cologne-1	AML_0005	M	Gastric juice	Neonatal ICU	2014-11-27	0
0003	Cologne-1	AML_0001	F	Pleural aspirate	Neonatal ICU	2014-11-25	0
	Singleton-2	AML_0029	M	Rectal swab	Pediatric general	2015-08-20	0
	Singleton-1	AML_0406		Siphon (environmental)	Neonatal ward	2017-11-01	
0002	Cologne-2	AML_0214	M	Rectal swab	Neonatal ICU	2017-01-12	0
0002	Cologne-2	AML_0213	M	Rectal swab	Neonatal ICU	2017-01-16	0
0002	Cologne-2	AML_0216		Weighing scale (environmental)	Neonatal ICU	2017-01-20	
0002	Cologne-2	AML_0217		Kleenex box (environmental)	Neonatal ICU	2017-01-20	
0002	Cologne-2	AML_0215	M	Ear swab	Neonatal ward	2017-01-16	0
0004	Cologne-4	AML_0027	M	Rectal swab	Pediatric general	2015-08-10	0
0004	Cologne-4	AML_0028	F	Gastric juice	Pediatric general	2015-08-13	0
0005	Cologne-3	AML_0293	M	Nose/Throat swab	Neonatal ICU	2017-12-12	0
0005	Cologne-3	AML_0294	F	Rectal swab	Neonatal ward	2017-12-19	0

Figure 3



Cluster	Patient Info	Isolate ID	Gender	Type of Specimen	Ward	Clinical Presentation	Isolation Date	Age (years)
0001		Smarc00428	M	Blood culture	Pediatric general	bacteraemia	2015-09-16	4
0001		Smarc00429	M	Nose/Throat swab	Emergency	infection	2015-10-12	0
		Smarc00478	M	blood culture	Pediatric general	bacteraemia	2015-12-13	12
0002		Smarc00449	M	Anal swab	Neonatal ICU	colonisation	2015-11-02	0
0002		Smarc00455		Thermometer (environmental)	Neonatal ICU	hospital surface	unknown	
0002	patient_D	Smarc00438	M	Anal swab	Neonatal ICU	colonisation	2015-10-19	0
0002		Smarc00466	M	Anal swab	Pediatric general	colonisation	2015-11-20	0
0002	patient_A	Smarc00442	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_I	Smarc00477	F	Wound swab	Neonatal ICU	colonisation	2015-12-09	0
0002		Smarc00445	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-10-22	0
0002	patient_E	Smarc00452	M	Nose/Throat swab	Neonatal ICU	bacteraemia	2015-11-02	0
0002	patient_D	Smarc00457	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-10-26	0
0002	patient_F	Smarc00462	F	Anal swab	Pediatric general	colonisation	2015-11-17	0
0002	patient_A	Smarc00431	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-10-12	0
0002		Smarc00432	unknown	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_B	Smarc00433	F	Blood culture	Neonatal ICU	bacteraemia	2015-10-16	0
0002	patient_B	Smarc00434	F	Nose/Throat swab	Neonatal ICU	bacteraemia	2015-10-19	0
0002		Smarc00435	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_C	Smarc00437	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_A	Smarc00443-A	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_A	Smarc00443-B	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_A	Smarc00444	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002		Smarc00446	F	Nose/Throat swab	Pediatric general	colonisation	2015-10-21	1
0002		Smarc00448		Thermometer (environmental)	Neonatal ICU	hospital surface	2015-10-21	
0002	patient_E	Smarc00450	M	Secretion	Neonatal ICU	bacteraemia	2015-11-02	0
0002		Smarc00456		Milk pump (environmental)	Neonatal ICU	hospital surface	unknown	
0002	patient_D	Smarc00458	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-10-26	0
0002		Smarc00459	F	Anal swab	Neonatal ICU	colonisation	2015-11-04	0
0002		Smarc00460	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-11-05	0
0002	patient_G	Smarc00464	F	Anal swab	Pediatric general	colonisation	2015-11-20	0
0002	patient_G	Smarc00465	F	Anal swab	Pediatric general	colonisation	2015-11-20	0
0002	patient_H	Smarc00468	M	Anal swab	Pediatric general	infection	2015-11-23	0
0002	patient_D	Smarc00469	M	Liquid other	Neonatal ICU	colonisation	2015-11-23	0
0002	patient_I	Smarc00476	F	Nose/Throat swab	Neonatal ICU	colonisation	2015-12-09	0
0002	patient_B	Smarc00436	F	Anal swab	Neonatal ICU	bacteraemia	2015-10-19	0
0002	patient_E	Smarc00451	M	Anal swab	Neonatal ICU	bacteraemia	2015-11-02	0
0002		Smarc00461	M	Anal swab	Neonatal ICU	colonisation	2015-11-06	0
0002	patient_C	Smarc00441	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002		Smarc00454		Thermometer (environmental)	Neonatal ICU	hospital surface	unknown	
0002	patient_C	Smarc00439-A	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_C	Smarc00440-S12	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_A	Smarc00481	M	Swab	Neonatal ICU	colonisation	2015-11-30	0
0002	patient_A	Smarc00475	M	Nose/Throat swab	Neonatal ICU	colonisation	2015-11-30	0
0002	patient_C	Smarc00439-B	M	Anal swab	Neonatal ICU	colonisation	2015-10-18	0
0002	patient_F	Smarc00463	F	Anal swab	Pediatric general	colonisation	2015-11-17	0
0002		Smarc00479		Thermometer (environmental)	Neonatal ICU	hospital surface	2015-12-10	
0002		Smarc00480		Thermometer (environmental)	Neonatal ICU	hospital surface	2015-12-10	
0002		Smarc00430	F	Swab	Emergency	colonisation	2015-10-12	0
0002		Smarc00453		Thermometer (environmental)	Neonatal ICU	hospital surface	unknown	
0002	patient_H	Smarc00474	M	Wound swab	Neonatal ICU	infection	2015-11-30	0
		Smarc00447	M	Nose/Throat swab	Pediatric general	unknown	2015-10-16	0

Figure 4A

Figure 4B

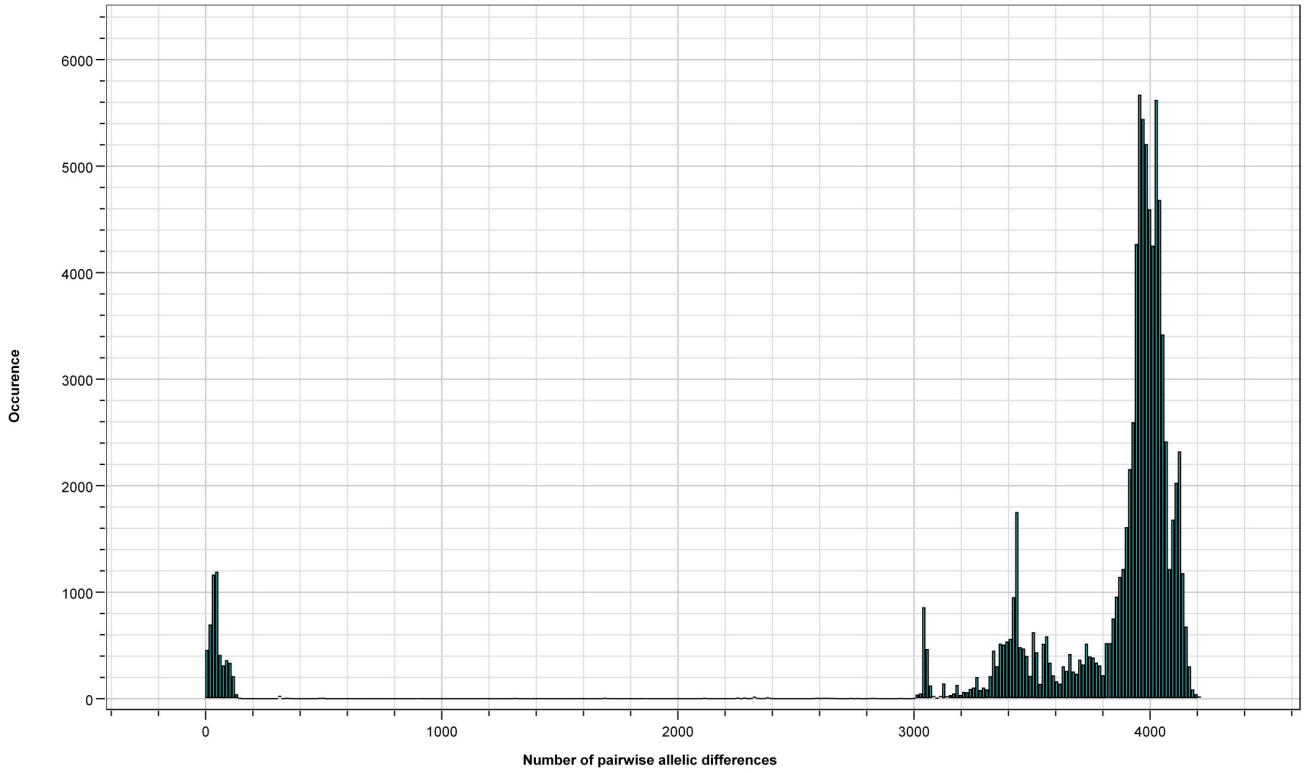
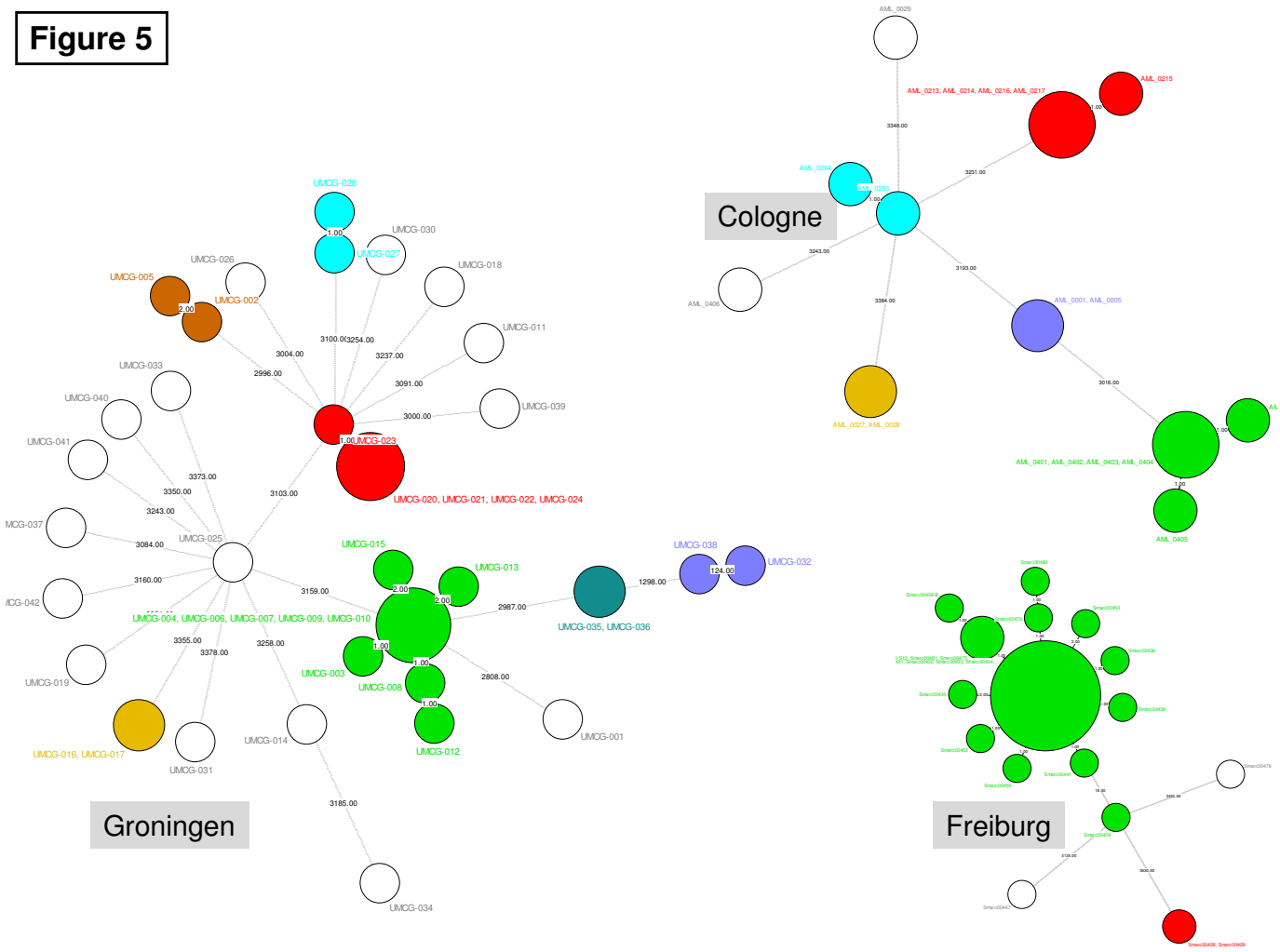


Figure 4C

Figure 5



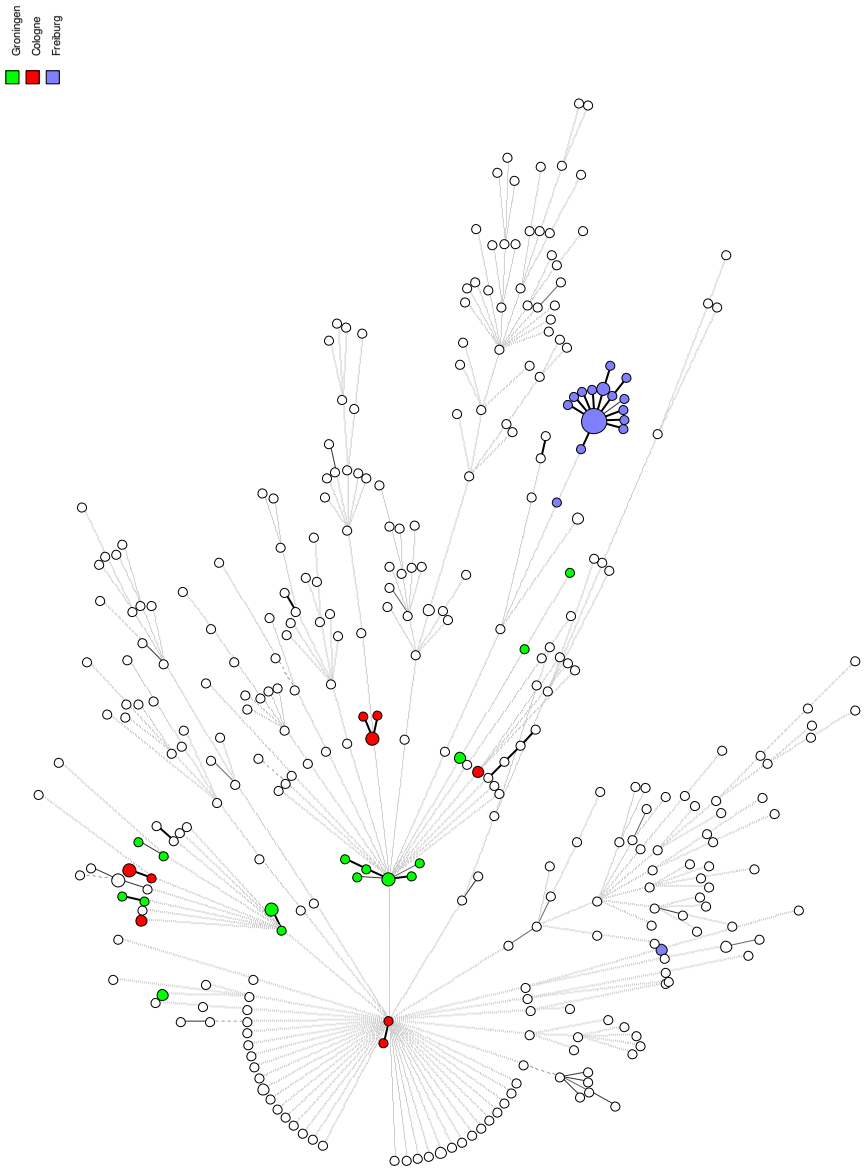


Figure 6